Advanced Algorithms

(Martingales and the Method of Bounded Differences)

(Some) Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$,
let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$, then:

for any $\delta > 0$,

$$\mathbb{P}(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

for $0 < \delta < 1$,

$$\mathbb{P}(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let $X = \sum_{i=1}^{n} X_i$, then:

for any $t > 0$,

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\right)$$

(Some) Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0,1\}$, let $X = \sum_{i=1}^n X_i$, and $\mu = \mathbb{E}(X)$, then:

for any $\delta > 0$,

$$\mathbb{P}(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

for $0 < \delta < 1$,

$$\mathbb{P}(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$, let $X = \sum_{i=1}^n X_i$, then:

for any $t > 0$,

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$\mathbb{P}(\mathcal{E}_1|\mathcal{E}_2) = \frac{\mathbb{P}(\mathcal{E}_1 \wedge \mathcal{E}_2)}{\mathbb{P}(\mathcal{E}_2)}$$

Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$\mathbb{P}(\mathcal{E}_1|\mathcal{E}_2) = \frac{\mathbb{P}(\mathcal{E}_1 \wedge \mathcal{E}_2)}{\mathbb{P}(\mathcal{E}_2)}$$

Example:
roll a fair six-sided dice
$\mathcal{E}_1$ = the outcome is six
$\mathcal{E}_2$ = the outcome is an even number

Conditional Probability

The *conditional probability* that event $\mathcal{E}_1$ occurs, given that event $\mathcal{E}_2$ occurs, is:

$$\mathbb{P}(\mathcal{E}_1|\mathcal{E}_2) = \frac{\mathbb{P}(\mathcal{E}_1 \wedge \mathcal{E}_2)}{\mathbb{P}(\mathcal{E}_2)}$$

Example:
roll a fair six-sided dice
$\mathcal{E}_1$ = the outcome is six
$\mathcal{E}_2$ = the outcome is an even number

$$\mathbb{P}(\mathcal{E}_1|\mathcal{E}_2) = \frac{1/6}{1/2} = \frac{1}{3}$$

Conditional Expectation

The *conditional expectation* of a random variable Y with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)$$

Conditional Expectation

The *conditional expectation* of a random variable Y with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)$$

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

Conditional Expectation

The *conditional expectation* of a random variable Y with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)$$

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y) = ?$

Conditional Expectation

The *conditional expectation* of a random variable Y with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)$$

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y) = ?$

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

Conditional Expectation

The *conditional expectation* of a random variable Y with respect to an event $\mathcal{E}$ is:

$$\mathbb{E}(Y \mid \mathcal{E}) = \sum_y y \cdot \mathbb{P}(Y = y \mid \mathcal{E})$$

In particular, if the event $\mathcal{E}$ is $X = x$, then:

$$\mathbb{E}(Y \mid X = x) = \sum_y y \cdot \mathbb{P}(Y = y \mid X = x)$$

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y) =?$

$\mathbb{E}(Y \mid X = \text{"China"}) =?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) =?$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$$\mathbb{E}(Y \mid X = \text{"China"}) = ?$$

$$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$\mathbb{E}(Y \mid X = x)$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(x) = \mathbb{E}(Y \mid X = x)$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \mathbb{E}(Y \mid X)$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) =?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) =?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$\mathbb{E}(X_1 \mid X_2 = a) =$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$$\mathbb{E}(Y \mid X = \text{"China"}) = ?$$

$$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$$

$$\mathbb{E}(X_1 \mid X_2) =$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$

$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$

$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$

$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) =$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$\mathbb{E}(Y \mid X = \text{"China"}) = ?$

$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$$

$$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$$

$$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n - a - b)/4$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$$\mathbb{E}(Y \mid X = \text{"China"}) = ?$$

$$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$$\mathbb{E}(X_1 \mid X_2 = a) = (n-a)/5$$

$$\mathbb{E}(X_1 \mid X_2) = (n-X_2)/5$$

$$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n-a-b)/4$$

$$\mathbb{E}(X_1 \mid X_2, X_3) =$$

Conditional Expectation

Example:
sample a human being uniformly at random
Y: height of the chosen human being
X: country of origin of the chosen human being

$$\mathbb{E}(Y \mid X = \text{"China"}) = ?$$

$$\mathbb{E}(Y \mid X = \text{"U.S."}) = ?$$

$$f(X) = \boxed{\mathbb{E}(Y \mid X)}$$

a random variable

Example:
throw a fair six-sided dice for n times
X_i: # of times i appears in n throws

$$\mathbb{E}(X_1 \mid X_2 = a) = (n - a)/5$$

$$\mathbb{E}(X_1 \mid X_2) = (n - X_2)/5$$

$$\mathbb{E}(X_1 \mid X_2 = a, X_3 = b) = (n - a - b)/4$$

$$\mathbb{E}(X_1 \mid X_2, X_3) = (n - X_2 - X_3)/4$$

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)$$

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)$$

average height of all human beings
= weighted average of the country-by-country average heights

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)$$

average height of all human beings
= weighted average of the country-by-country average heights

$$\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)$$

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)$$

average height of all human beings
= weighted average of the country-by-country average heights

$$\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)$$

average height of all male/female human beings
= weighted average of the country-by-country average male/female heights

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)$$

average height of all human beings
= weighted average of the country-by-country average heights

$$\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)$$

average height of all male/female human beings
= weighted average of the country-by-country average male/female heights

$$\mathbb{E}\big(\mathbb{E}(f(X)g(X,Y) \mid X)\big) = \mathbb{E}\big(f(X)\mathbb{E}(g(X,Y) \mid X)\big)$$

Fundamental Facts about Conditional Expectation

Example: Y: height of the chosen human being
X: country of origin of the chosen human being
Z: gender of the chosen human being

$$\boxed{\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)}$$

average height of all human beings
= weighted average of the country-by-country average heights

$$\boxed{\mathbb{E}(Y \mid Z) = \mathbb{E}(\mathbb{E}(Y \mid X, Z) \mid Z)}$$

average height of all male/female human beings
= weighted average of the country-by-country average male/female heights

$$\boxed{\mathbb{E}\big(\mathbb{E}(f(X)g(X,Y) \mid X)\big) = \mathbb{E}\big(f(X)\mathbb{E}(g(X,Y) \mid X)\big)}$$

once X is fixed to some x,
$\mathbb{E}(f(X)g(X,Y) \mid X = x) = f(x)\mathbb{E}(g(X,Y) \mid X = x)$

Martingales

originally refers to a betting strategy:

“double your bet after every loss”

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^i = 1$

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^{i} = 1$

consider a fair game, with any betting strategy

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^i = 1$

consider a fair game, with any betting strategy

let X_i be our wealth after i rounds

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^{i} = 1$

consider a fair game, with any betting strategy

let X_i be our wealth after i rounds

$$\mathbb{E}(X_{i+1} \mid X_0, X_1, \cdots, X_i) =$$

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^i = 1$

consider a fair game, with any betting strategy

let X_i be our wealth after i rounds

$$\mathbb{E}(X_{i+1} \mid X_0, X_1, \cdots, X_i) = X_i$$

Martingales

originally refers to a betting strategy:

"double your bet after every loss"

when you get a win after n losses: $2^{\mathrm{n}} - \sum_{i=0}^{n-1} 2^{i} = 1$

consider a fair game, with any betting strategy

let X_i be our wealth after i rounds

$$\mathbb{E}(X_{i+1} \mid X_0, X_1, \cdots, X_i) = X_i$$

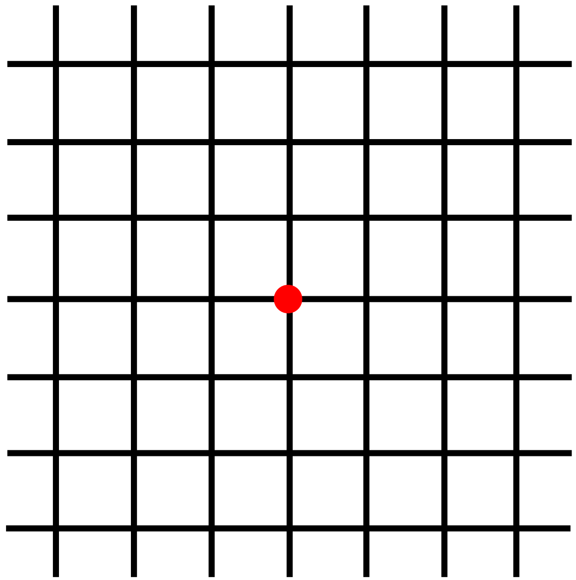
since the game is **fair**, conditioned on past history,
we expect **no change** to current value after one round

Martingales

A sequence of random variables $X_0, X_1, \cdots$ is *a martingale* if for all $i \geq 1$,

$$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$$

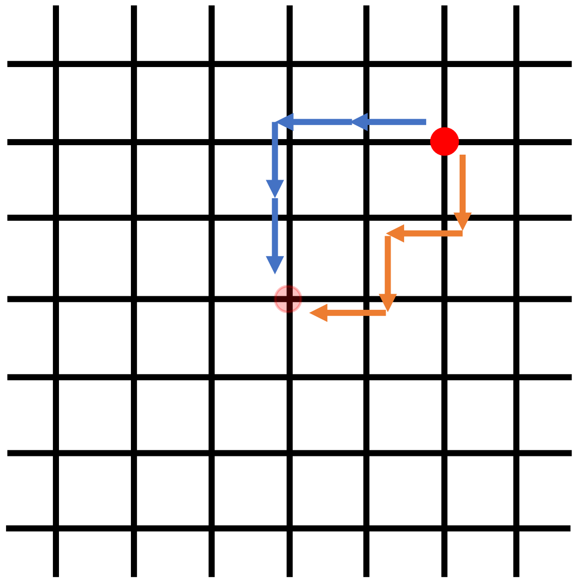
Example: Random Walk



a dot starting from the origin

in each step, move equiprobably to one of four neighbor

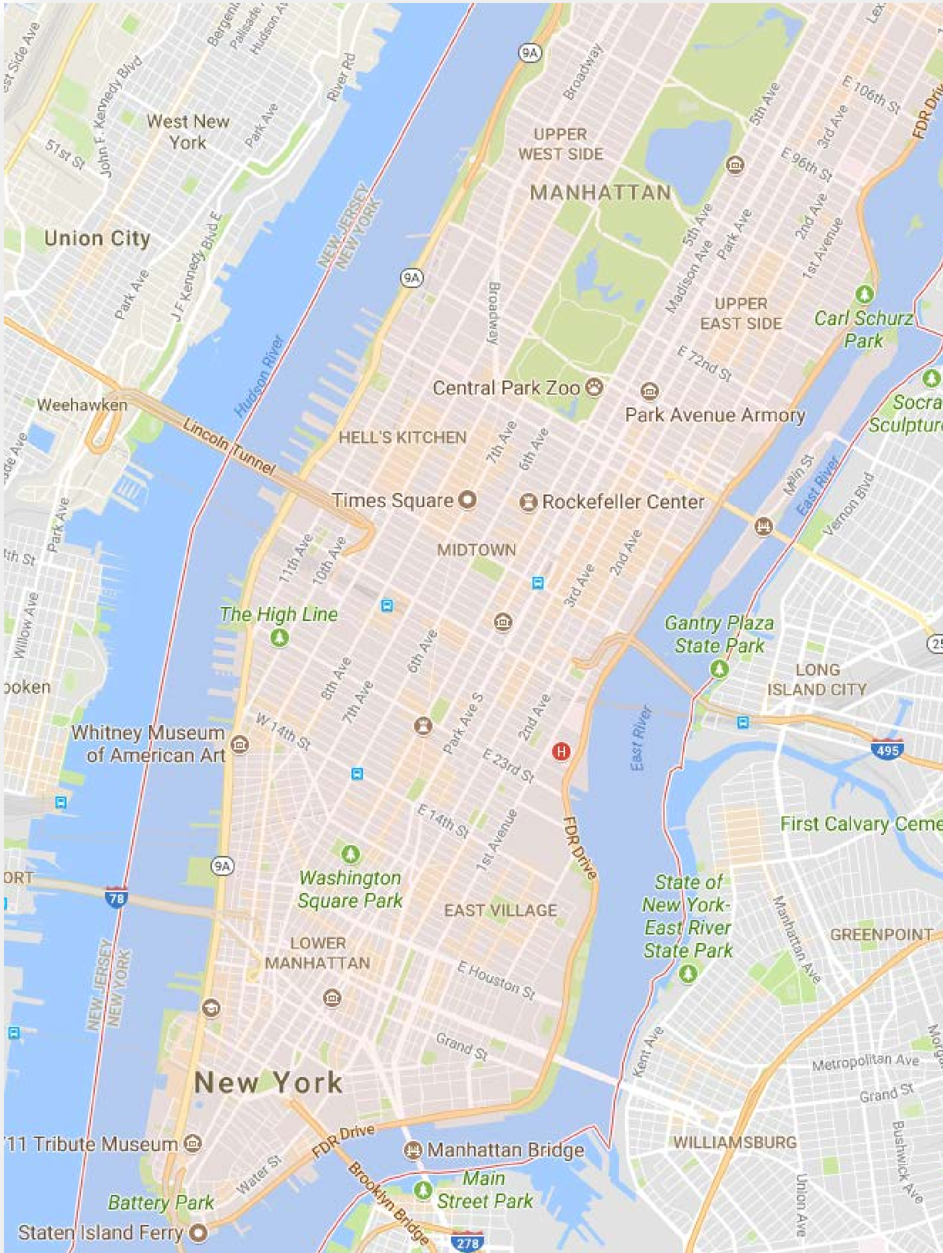
Example: Random Walk



a dot starting from the origin

in each step, move equiprobably to one of four neighbor

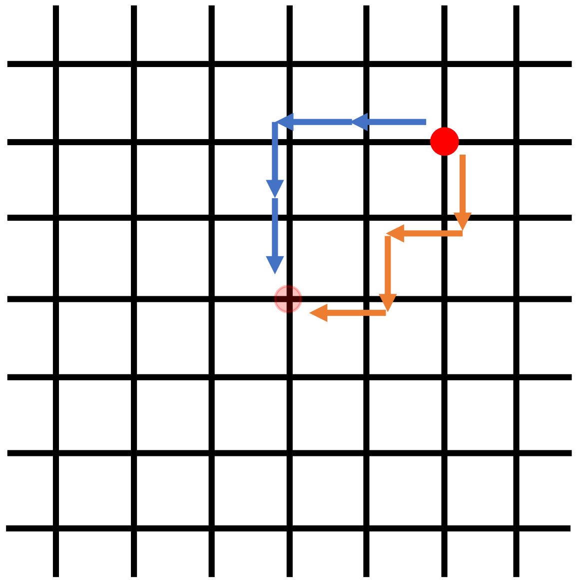
after i steps, use X_i to denote # of hops to origin (Manhattan distance)



dom Walk

- ot starting from the origin
- ch step, move equiprobably to one of four neighbor
- er i steps, use X_i to denote s to origin (Manhattan distance)

Example: Random Walk



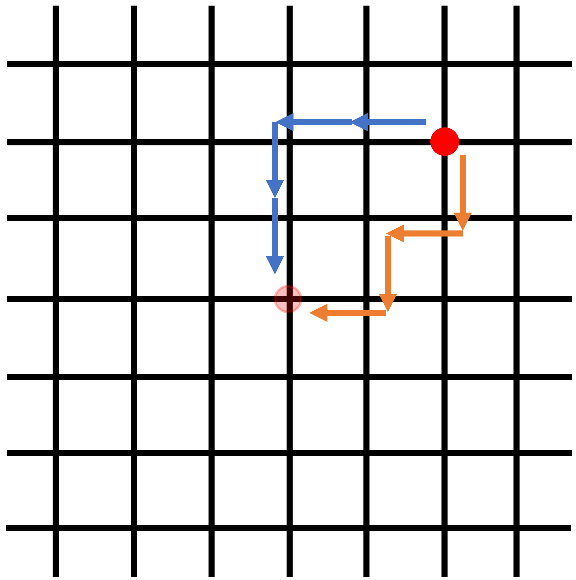
a dot starting from the origin

in each step, move equiprobably to one of four neighbor

after i steps, use X_i to denote # of hops to origin (Manhattan distance)

$$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$$

Example: Random Walk



a dot starting from the origin

in each step, move equiprobably to one of four neighbor

after i steps, use X_i to denote # of hops to origin (Manhattan distance)

$$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$$

How far the dot is away from the origin after n steps?

Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

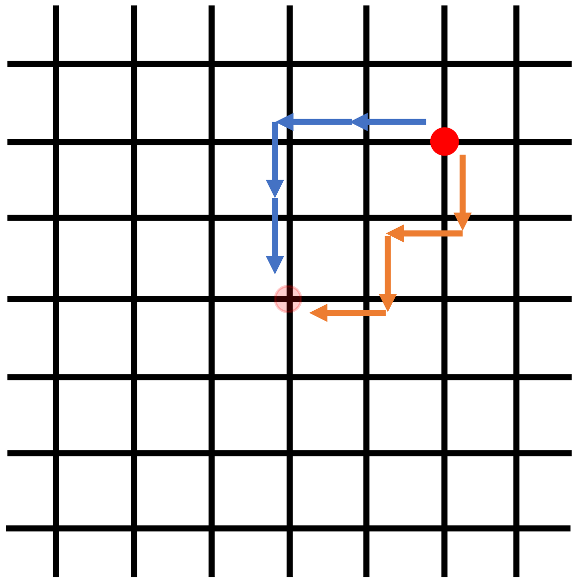
$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

$X_0, X_1, \cdots$ are not necessarily independent

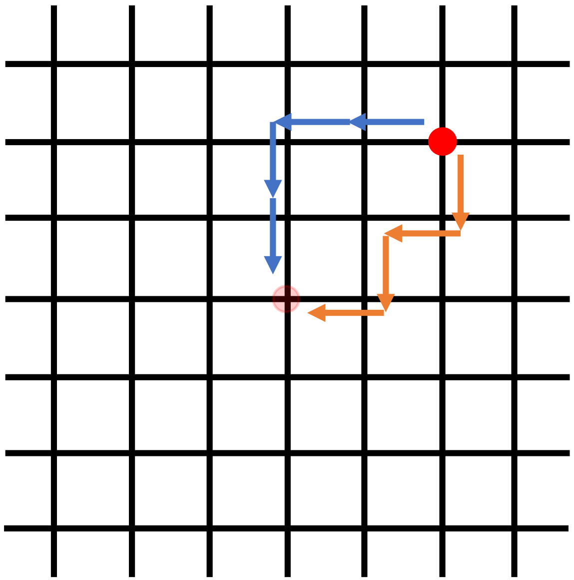
Azuma's Inequality in Action



After i steps, use X_i to denote
# of hops to origin (Manhattan distance)

How large is X_n?

Azuma's Inequality in Action



After i steps, use X_i to denote # of hops to origin (Manhattan distance)

How large is X_n?

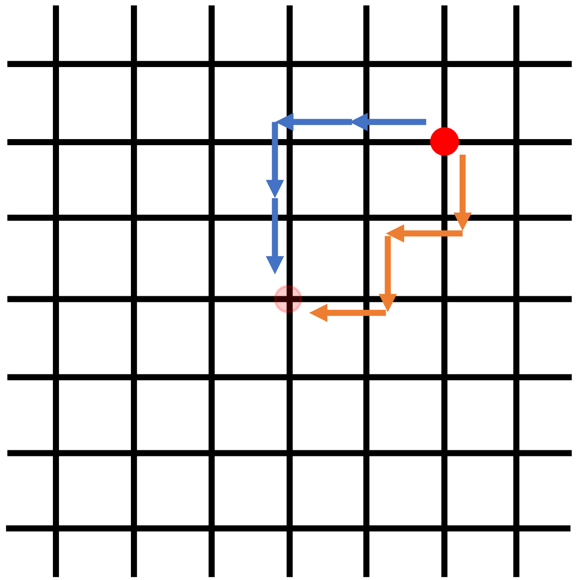
Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

Azuma's Inequality in Action



After i steps, use X_i to denote
# of hops to origin (Manhattan distance)

How large is X_n?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

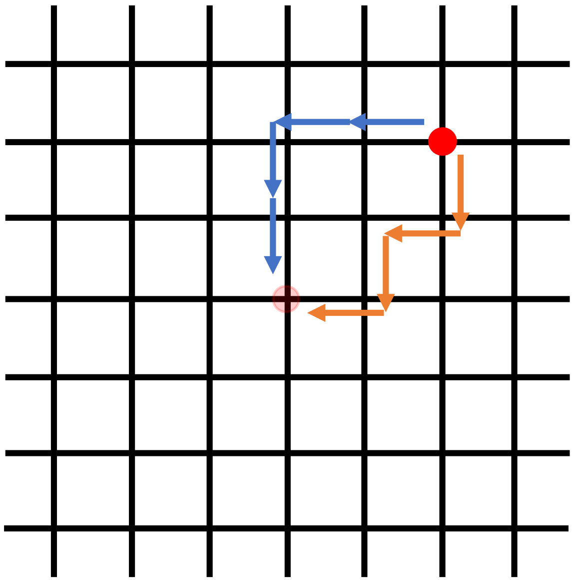
$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

We know $X_0 = 0$, and $|X_k - X_{k-1}| \leq 1$

Azuma's Inequality in Action



After i steps, use X_i to denote # of hops to origin (Manhattan distance)

How large is X_n?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

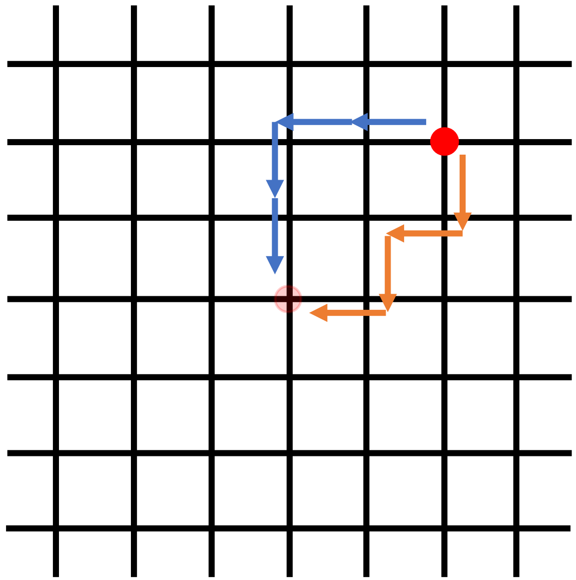
Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

We know $X_0 = 0$, and $|X_k - X_{k-1}| \leq 1$

$$\mathbb{P}(|X_n| \geq c\sqrt{n}) \leq 2e^{-c^2/2}$$

Azuma's Inequality in Action



After i steps, use X_i to denote
# of hops to origin (Manhattan distance)

How large is X_n?

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

We know $X_0 = 0$, and $|X_k - X_{k-1}| \leq 1$

$$\mathbb{P}(|X_n| \geq c\sqrt{n}) \leq 2e^{-c^2/2}$$

Within $O(\sqrt{n \log n})$ w.h.p.

Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

For a sequence of r.v., if in each step:

* on average make no change to current value (martingale)
* no big jump (bounded difference)

Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

For a sequence of r.v., if in each step:

* on average make no change to current value (martingale)
* no big jump (bounded difference)

Then final value does not deviate far from the initial.

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Use similar strategy as in proving Chernoff bound:

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Use similar strategy as in proving Chernoff bound:

(a) Apply generalized Markov's inequality to MGF

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Use similar strategy as in proving Chernoff bound:

(a) Apply generalized Markov's inequality to MGF

(b)* Bound the value of MGF (use Hoeffding's lemma)

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|X_n - X_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

Use similar strategy as in proving Chernoff bound:

(a) Apply generalized Markov's inequality to MGF

(b)* Bound the value of MGF (use Hoeffding's lemma)

(c) Optimize the value of MGF

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\boxed{\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)}$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\boxed{\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)}$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

W.l.o.g., assume $X_0 = 0$. (Otherwise, do the proof w.r.t. sequence $X_i' = X_i - X_0$.)

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\boxed{\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)}$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

W.l.o.g., assume $X_0 = 0$. (Otherwise, do the proof w.r.t. sequence $X_i' = X_i - X_0$.)

Define $Y_i = X_i - X_{i-1}$, it is easy to see $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$.

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i|\boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\mathbb{E}(e^{\lambda X_n})}{e^{\lambda t}}$$

for $\lambda > 0$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$ $\qquad\qquad$ $\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right)$$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$\boxed{\mathbb{E}(Y) = \mathbb{E}\big(\mathbb{E}(Y \mid X)\big)}$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$\boxed{\quad\quad\quad} = \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})\right)$$

$$\boxed{\mathbb{E}\big(\mathbb{E}(f(X) g(X,Y) | X)\big) = \mathbb{E}\big(f(X) \mathbb{E}(g(X,Y) \mid X)\big)}$$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$ $\qquad$ $\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$\boxed{\phantom{\mathbb{E}(e^{\lambda X_n})}} = \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$\boxed{\phantom{\mathbb{E}(e^{\lambda X_n})}} = \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \le \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \le \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$\boxed{\quad\quad} = \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right)$$

$$\boxed{\quad\quad} \leq \exp\left(\frac{\lambda^2 4 c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \le c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \ge t) \le ?$

$$\mathbb{P}(X_n - X_0 \ge t) = \mathbb{P}(X_n \ge t) = \mathbb{P}(e^{\lambda X_n} \ge e^{\lambda t}) \le \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right) \le \mathbb{E}\left(e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2}\right) = e^{\lambda^2 c_n^2/2}\, \mathbb{E}\left(e^{\lambda X_{n-1}}\right)$$

$$\le \exp\left(\frac{\lambda^2 4 c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \le \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$

$\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right) \leq \mathbb{E}\left(e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2}\right) = e^{\lambda^2 c_n^2/2} \mathbb{E}\left(e^{\lambda X_{n-1}}\right)$$

$$\leq e^{\lambda^2 c_n^2/2} e^{\lambda^2 c_{n-1}^2/2} \mathbb{E}(e^{\lambda X_{n-2}}) \leq \cdots \leq \left(\prod_{k=1}^{n} e^{\lambda^2 c_k^2/2}\right) \mathbb{E}(e^{\lambda X_0}) = e^{(\lambda^2/2)\sum_{k=1}^{n} c_k^2}$$

$$\leq \exp\left(\frac{\lambda^2 4 c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

$X_0, X_1, \cdots$ is a martingale, with $X_0 = 0$, and $|X_k - X_{k-1}| \leq c_k$

$Y_i = X_i - X_{i-1}$, thus $\mathbb{E}(Y_i | \boldsymbol{X}_{i-1}) = 0$ $\qquad\qquad$ $\mathbb{P}(X_n - X_0 \geq t) \leq ?$

$$\mathbb{P}(X_n - X_0 \geq t) = \mathbb{P}(X_n \geq t) = \mathbb{P}(e^{\lambda X_n} \geq e^{\lambda t}) \leq \frac{\boxed{\mathbb{E}(e^{\lambda X_n})}}{e^{\lambda t}}$$

for $\lambda > 0$

$$\leq \frac{e^{(\lambda^2/2)\sum_{k=1}^n c_k^2}}{e^{\lambda t}} = \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

minimized when $\lambda = \frac{t}{\sum_{k=1}^n c_k^2}$

$$= \mathbb{E}\left(e^{\lambda(Y_n + X_{n-1})}\right) = \mathbb{E}\left(e^{\lambda Y_n} e^{\lambda X_{n-1}}\right) = \mathbb{E}\left(\mathbb{E}(e^{\lambda Y_n} e^{\lambda X_{n-1}} \mid \boldsymbol{X_{n-1}})\right)$$

$$= \mathbb{E}\left(e^{\lambda X_{n-1}} \boxed{\mathbb{E}(e^{\lambda Y_n} \mid \boldsymbol{X_{n-1}})}\right) \leq \mathbb{E}\left(e^{\lambda X_{n-1}} e^{\lambda^2 c_n^2/2}\right) = e^{\lambda^2 c_n^2/2} \mathbb{E}\left(e^{\lambda X_{n-1}}\right)$$

$$\leq e^{\lambda^2 c_n^2/2} e^{\lambda^2 c_{n-1}^2/2} \mathbb{E}(e^{\lambda X_{n-2}}) \leq \cdots \leq \left(\prod_{k=1}^n e^{\lambda^2 c_k^2/2}\right) \mathbb{E}(e^{\lambda X_0}) = e^{(\lambda^2/2)\sum_{k=1}^n c_k^2}$$

$$\leq \exp\left(\frac{\lambda^2 4 c_n^2}{8}\right) = \exp\left(\frac{\lambda c_n^2}{2}\right)$$

For any random variable $Z \in [a, b]$ with $\mathbb{E}(Z) = 0$,

$$\mathbb{E}(e^{\lambda Z}) \leq \exp\left(\frac{\lambda^2 (b-a)^2}{8}\right)$$

let $Z_n = (Y_n \mid \boldsymbol{X}_{n-1})$

$\mathbb{E}(Z_n) = \mathbb{E}(Y_n \mid \boldsymbol{X}_{n-1}) = 0$

$Z_n \in [-c_n, c_n]$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\boxed{\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)}$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right) \quad ???$$

Proving Azuma's Inequality

Let $X_0, X_1, \cdots$ be a martingale such that for all $k \geq 1$,

$$|X_k - X_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(X_n - X_0 \geq \quad t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

$$\mathbb{P}(X_n - X_0 \leq -t) \leq \exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$???

let $X_i' = -X_i$

$$\mathbb{P}(X_n - X_0 \leq -t) = \mathbb{P}((-X_i') - (-X_0') \leq -t) = \mathbb{P}(X_i' - X_0' \geq t) \leq \cdots$$

Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is *a martingale with respect to the sequence* $X_0, X_1, \cdots$ if for all $i \geq 0$,

- Y_i is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$

Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is *a martingale with respect to the sequence* $X_0, X_1, \cdots$ if for all $i \geq 0$,

- Y_i is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$

betting on a fair game

X_i: gain/loss of the i^{th} bet

Y_i: wealth after the i^{th} bet

Generalized Martingales

A sequence of random variables $Y_0, Y_1, \cdots$ is *a martingale with respect to the sequence* $X_0, X_1, \cdots$ if for all $i \geq 0$,

- Y_i is a function of $X_0, X_1, \cdots, X_i$
- $\mathbb{E}(Y_{i+1} \mid X_0, X_1, \cdots, X_i) = Y_i$

betting on a fair game

X_i: gain/loss of the i^{th} bet

Y_i: wealth after the i^{th} bet $\leftarrow$ martingale (since game is fair)

Generalized Azuma's Inequality

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Azuma's Inequality

martingale $X_0, X_1, \cdots$

with $|X_k - X_{k-1}| \le c_k$,

then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

$\downarrow$ generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

with $|Y_k - Y_{k-1}| \le c_k$,

then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

$\downarrow$ generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

$Y_i = f(X_0, X_1, \cdots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

Doob Sequence

The Doob sequence of a function f with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}\left(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i\right)$$

In particular,

$Y_0 = \mathbb{E}\left(f(X_1, \cdots, X_n)\right)$, and $Y_n = f(X_1, \cdots, X_n)$

Doob Sequence

$f($  ,  ,  ,  $)$

Doob Sequence

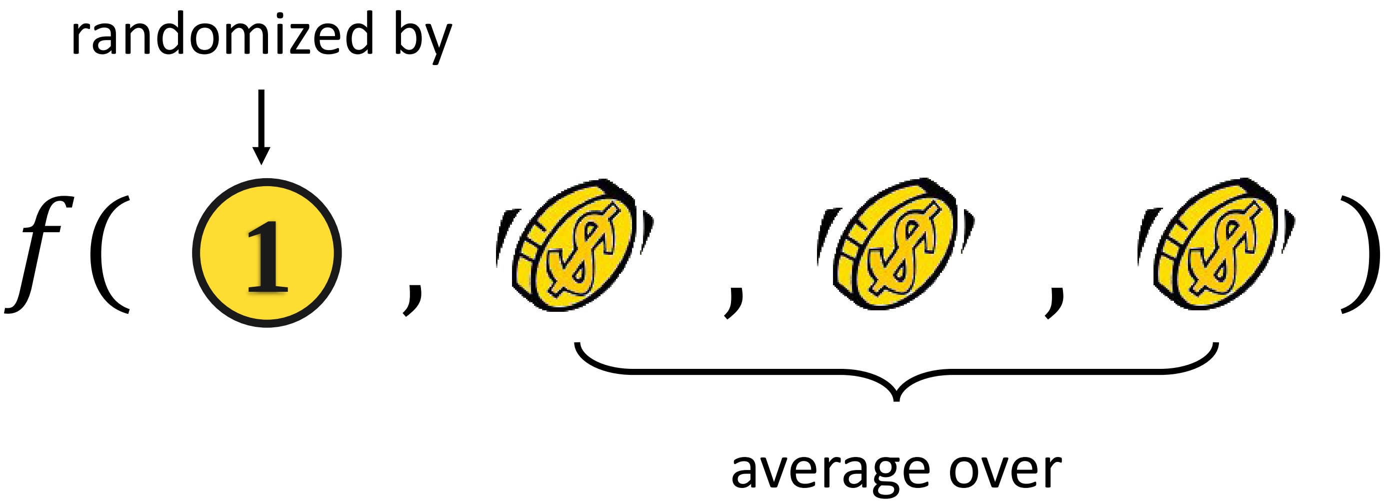
$f($  ,  ,  ,  $)$

average over

no information

$\mathbb{E}(f)$

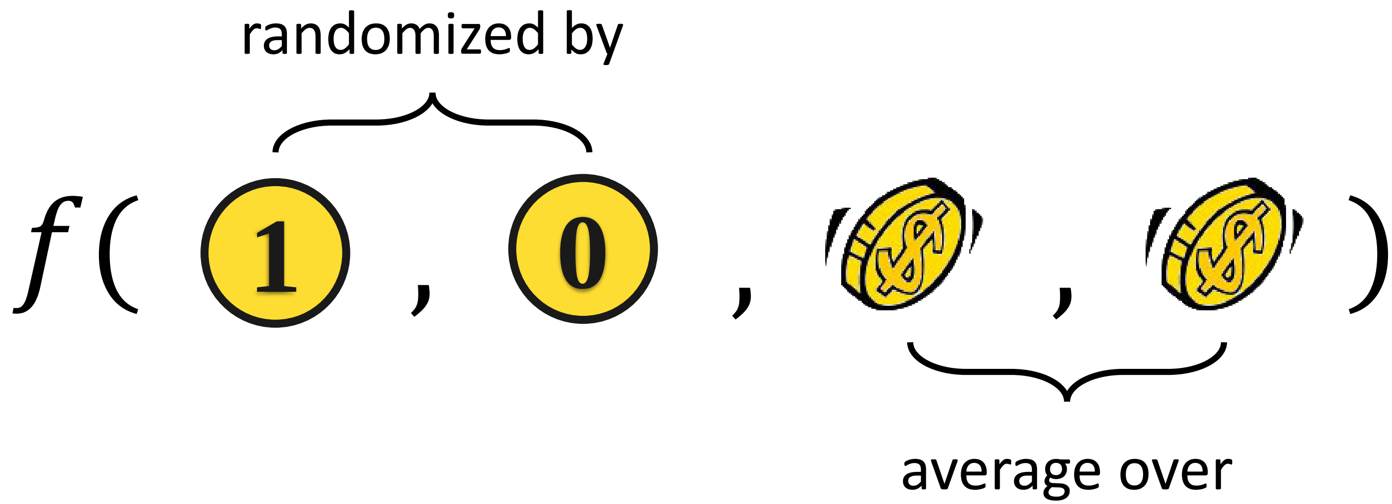
Doob Sequence



no information

$$\mathbb{E}(f) \longrightarrow \mathbb{E}(f \mid \boldsymbol{X}_1)$$

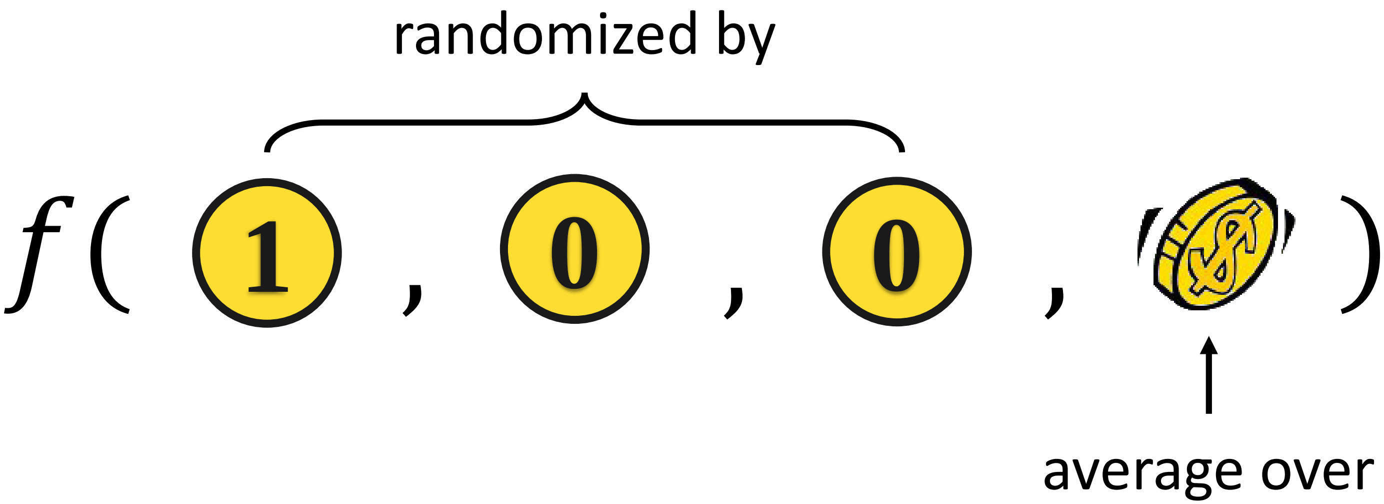
Doob Sequence



no information

$$\mathbb{E}(f) \longrightarrow \mathbb{E}(f\,|\,\boldsymbol{X}_1) \longrightarrow \mathbb{E}(f\,|\,\boldsymbol{X}_2)$$

Doob Sequence



no information

$$\mathbb{E}(f) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_1) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_2) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_3)$$

Doob Sequence

$$f(\overbrace{1, 0, 0, 1}^{\text{randomized by}})$$

no information

full information

$$\mathbb{E}(f) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_1) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_2) \longrightarrow \mathbb{E}(f|\boldsymbol{X}_3) \rightarrow \mathbb{E}(f|\boldsymbol{X}_4) = f(\boldsymbol{X}_4)$$

Doob Martingale

The Doob sequence of a function f with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}\left(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i\right)$$

The Doob sequence of a function f is a martingale. That is,

$$\mathbb{E}(Y_i | X_1, \cdots, X_{i-1}) = Y_{i-1}$$

Doob Martingale

The Doob sequence of a function f with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}\left(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i\right)$$

The Doob sequence of a function f is a martingale. That is,

$$\mathbb{E}(Y_i | X_1, \cdots, X_{i-1}) = Y_{i-1}$$

$$\mathbb{E}(Y_i \mid \boldsymbol{X}_{i-1}) = \mathbb{E}\left(\mathbb{E}(f(\boldsymbol{X}_n) \mid \boldsymbol{X}_i) \mid \boldsymbol{X}_{i-1}\right)$$

Doob Martingale

The Doob sequence of a function f with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

$$Y_i = \mathbb{E}\left(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i\right)$$

The Doob sequence of a function f is a martingale. That is,

$$\mathbb{E}(Y_i | X_1, \cdots, X_{i-1}) = Y_{i-1}$$

$$\begin{aligned}\mathbb{E}(Y_i \mid \boldsymbol{X}_{i-1}) &= \mathbb{E}\left(\mathbb{E}(f(\boldsymbol{X}_n) \mid \boldsymbol{X}_i) \mid \boldsymbol{X}_{i-1}\right) \\ &= \mathbb{E}(f(\boldsymbol{X}_n) \mid \boldsymbol{X}_{i-1})\end{aligned}$$

$$\boxed{\mathbb{E}(Y \mid \boldsymbol{Z}) = \mathbb{E}(\mathbb{E}(Y \mid \boldsymbol{X}, \boldsymbol{Z}) \mid \boldsymbol{Z})}$$

Doob Martingale

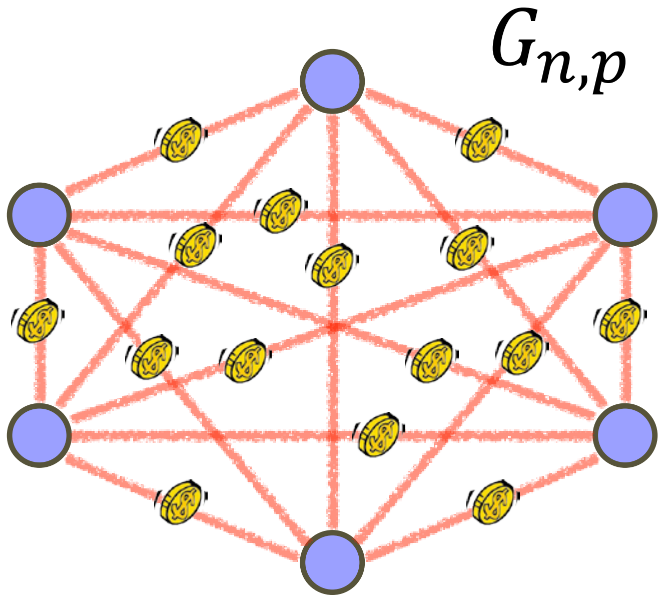
The Doob sequence of a function f with respect to a sequence of random variables $X_1, X_2, \cdots, X_n$ is

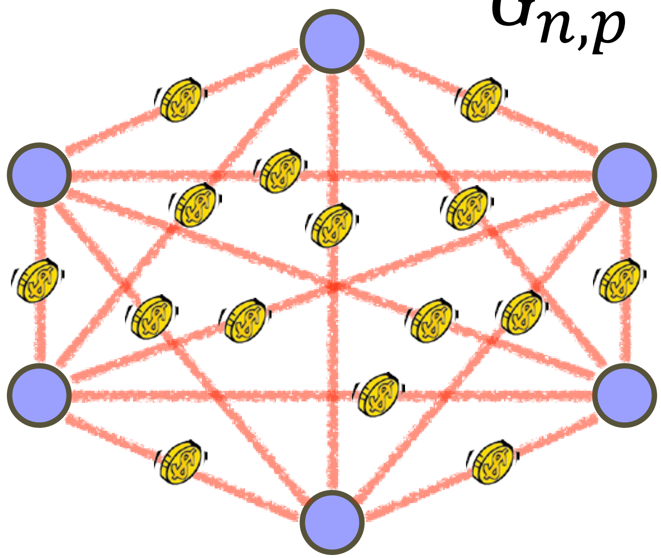
$$Y_i = \mathbb{E}\left(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i\right)$$

The Doob sequence of a function f is a martingale. That is,

$$\mathbb{E}(Y_i | X_1, \cdots, X_{i-1}) = Y_{i-1}$$

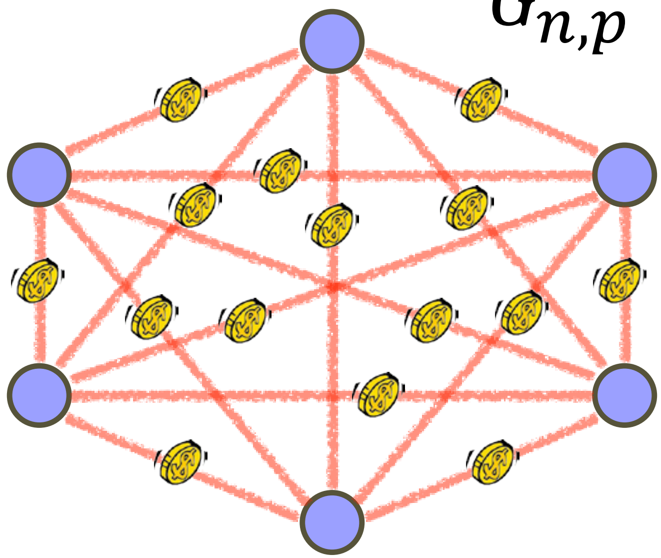
$$\begin{aligned}
\mathbb{E}(Y_i \mid \boldsymbol{X}_{i-1}) &= \mathbb{E}\left(\mathbb{E}(f(\boldsymbol{X}_n) \mid \boldsymbol{X}_i) \mid \boldsymbol{X}_{i-1}\right) \\
&= \mathbb{E}(f(\boldsymbol{X}_n) \mid \boldsymbol{X}_{i-1}) \\
&= Y_{i-1}
\end{aligned}$$


$G_{n,p}$



Graph parameter: $f(G)$

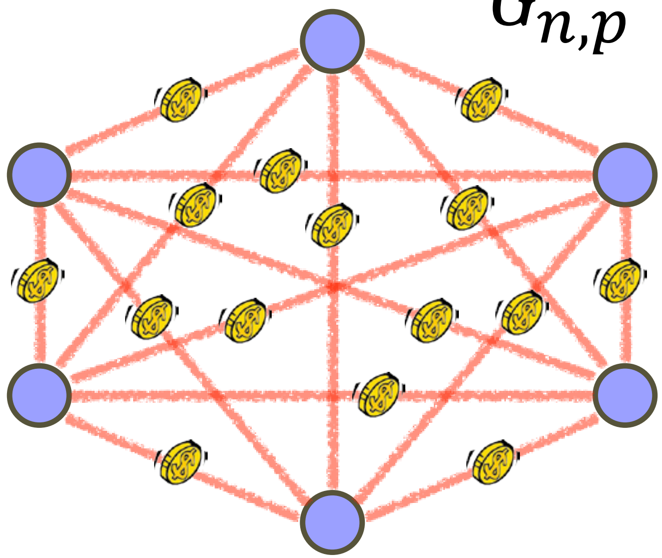
Example: components number,
chromatic number,
diameter



Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertex pairs: $1,2,3,\cdots,\binom{n}{2}$



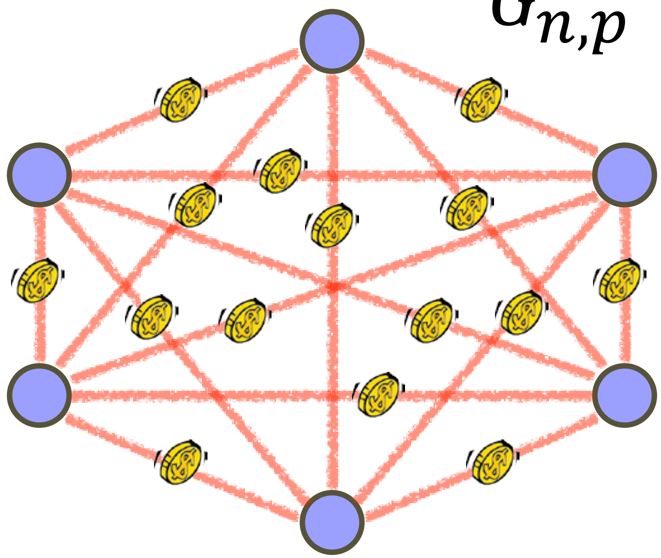
$G_{n,p}$

Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertex pairs: $1,2,3,\cdots,\binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases}$

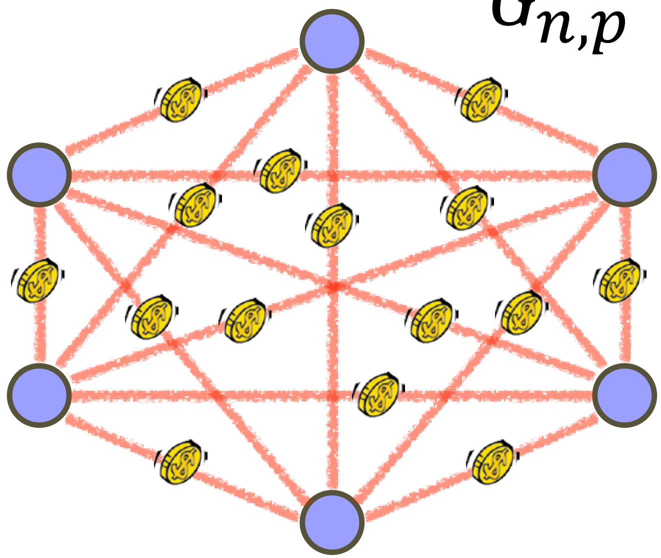


Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertex pairs: $1,2,3,\cdots,\binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases}$ $\quad Y_i = \mathbb{E}(f(G) \mid I_1, \cdots, I_i)$



Graph parameter: $f(G)$

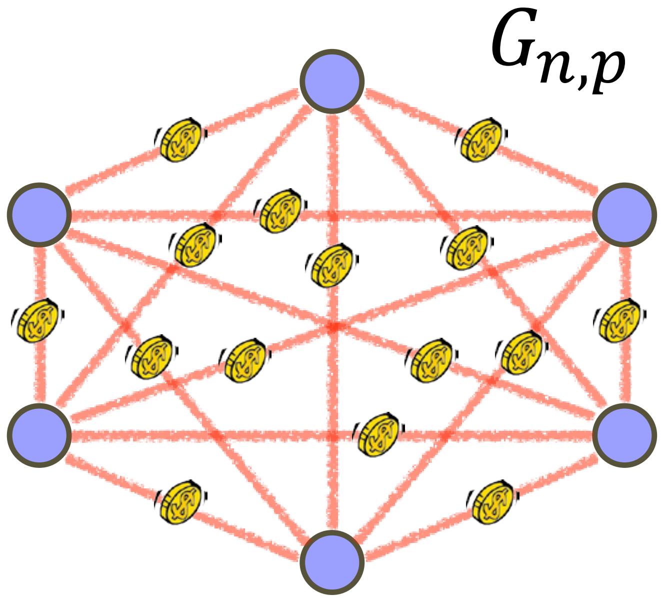
Example: components number,
chromatic number,
diameter

numbering all vertex pairs: $1,2,3,\cdots,\binom{n}{2}$

Define i.r.v. $I_j = \begin{cases} 1 & \text{edge } j \in G \\ 0 & \text{edge } j \notin G \end{cases}$ $\quad Y_i = \mathbb{E}(f(G) \mid I_1,\cdots,I_i)$

$Y_0, Y_1, \cdots, Y_{\binom{n}{2}}$ is a Doob sequence, called edge exposure martingale

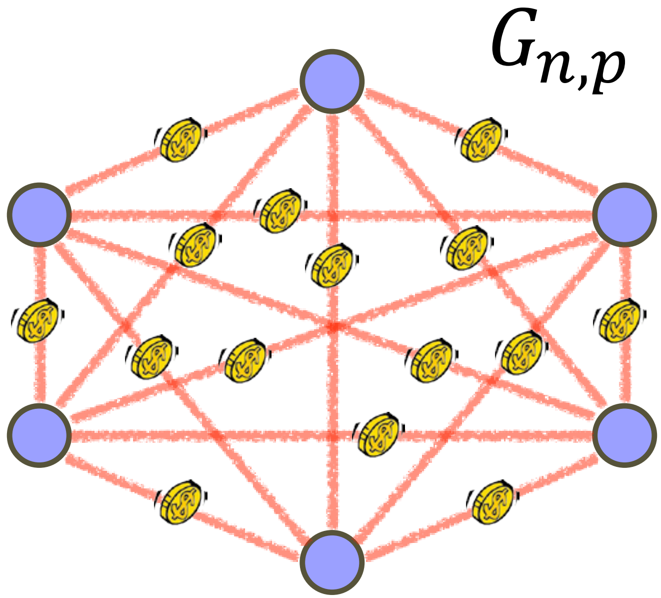
In particular, $Y_0 = \mathbb{E}(f(G))$, and $Y_{\binom{n}{2}} = f(G)$



Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertices: $1,2,3,\cdots,n$

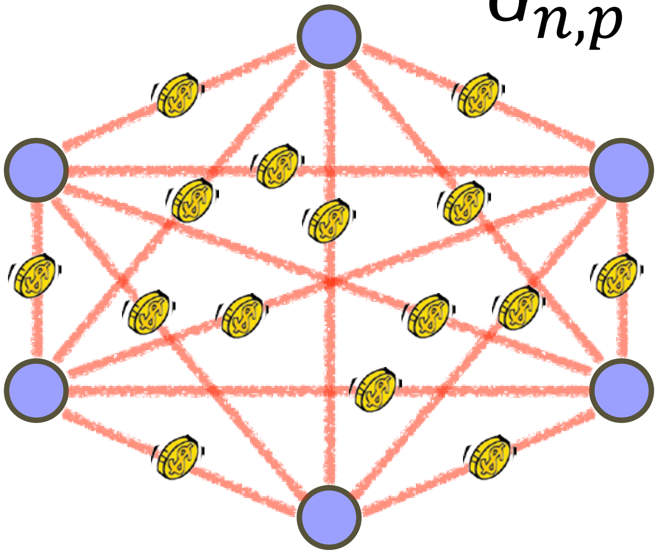


Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertices: $1,2,3,\cdots,n$

X_i: subgraph of G induced by the first i vertices



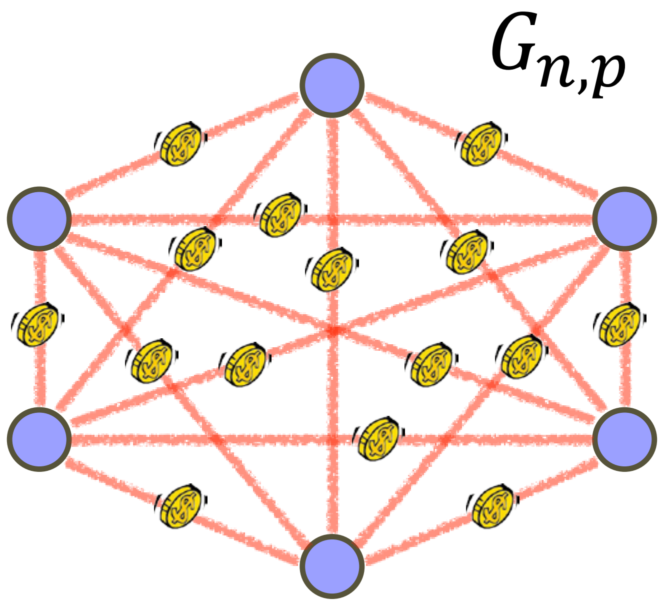
Graph parameter: $f(G)$

Example: components number,
chromatic number,
diameter

numbering all vertices: $1,2,3,\cdots,n$

X_i: subgraph of G induced by the first i vertices

$$Y_i = \mathbb{E}(f(G) \mid X_1, \cdots, X_i)$$



Graph parameter: $f(G)$

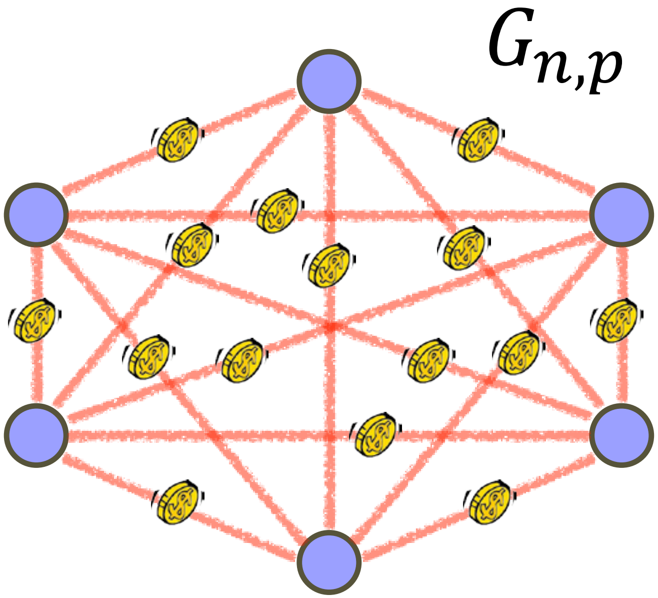
Example: components number,
chromatic number,
diameter

numbering all vertices: $1,2,3,\cdots,n$

X_i: subgraph of G induced by the first i vertices

$$Y_i = \mathbb{E}(f(G) \mid X_1, \cdots, X_i)$$

$Y_0, Y_1, \cdots, Y_n$ is a Doob sequence, called vertex exposure martingale
In particular, $Y_0 = \mathbb{E}(f(G))$, and $Y_n = f(G)$



chromatic number

$\chi(G)$

is the smallest number of colors to properly color G

numbering all vertices: $1,2,3,\cdots,n$

X_i: subgraph of G induced by the first i vertices

$$Y_i = \mathbb{E}(\chi(G) \mid X_1, \cdots, X_i)$$

$Y_0, Y_1, \cdots, Y_n$ is a Doob sequence (vertex exposure martingale)
In particular, $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

Generalized Azuma's Inequality in Action

Concentration of Chromatic Number

chromatic number $\chi(G)$

X_i: subgraph of G induced by the first i vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

Generalized Azu

Concentra

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

chromatic numb

X_i: subgraph of G induced by the first i vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

Generalized Azu

Concentra

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

chromatic numb

X_i: subgraph of G induced by the first i vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

A new vertex can always be given a new color!

Generalized Azu

Concentra

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^n c_k^2}\right)$$

chromatic numb

X_i: subgraph of G induced by the first i vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

A new vertex can always be given a new color!

$|Y_i - Y_{i-1}| \leq 1$

Generalized Az

Concentra

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$ such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

chromatic numb

X_i: subgraph of G induced by the first i vertices

$Y_i = \mathbb{E}(\chi(G) | X_1, \cdots, X_i)$

$Y_0, Y_1, \cdots, Y_n$ a Doob martingale: $Y_0 = \mathbb{E}(\chi(G))$, and $Y_n = \chi(G)$

A new vertex can always be given a new color!

$|Y_i - Y_{i-1}| \leq 1$

$$\mathbb{P}(|\chi(G) - \mathbb{E}(\chi(G))| \geq t\sqrt{n})$$

$$= \mathbb{P}(|Y_n - Y_0| \geq t\sqrt{n}) \leq 2e^{-t^2/2}$$

Tight Concentration of Chromatic Number

Theorem [Shamir & Spencer (1987)]:
Let $G \sim G(n,p)$, then:

$$\mathbb{P}(|\chi(G) - \mathbb{E}(\chi(G))| \geq t\sqrt{n}) \leq 2e^{-t^2/2}$$

Azuma's Inequality

martingale $X_0, X_1, \cdots$

with $|X_k - X_{k-1}| \le c_k$,

then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

with $|Y_k - Y_{k-1}| \le c_k$,

then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

$Y_i = f(X_0, X_1, \cdots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

Azuma's Inequality

martingale $X_0, X_1, \cdots$

with $|X_k - X_{k-1}| \le c_k$,

then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

with $|Y_k - Y_{k-1}| \le c_k$,

then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

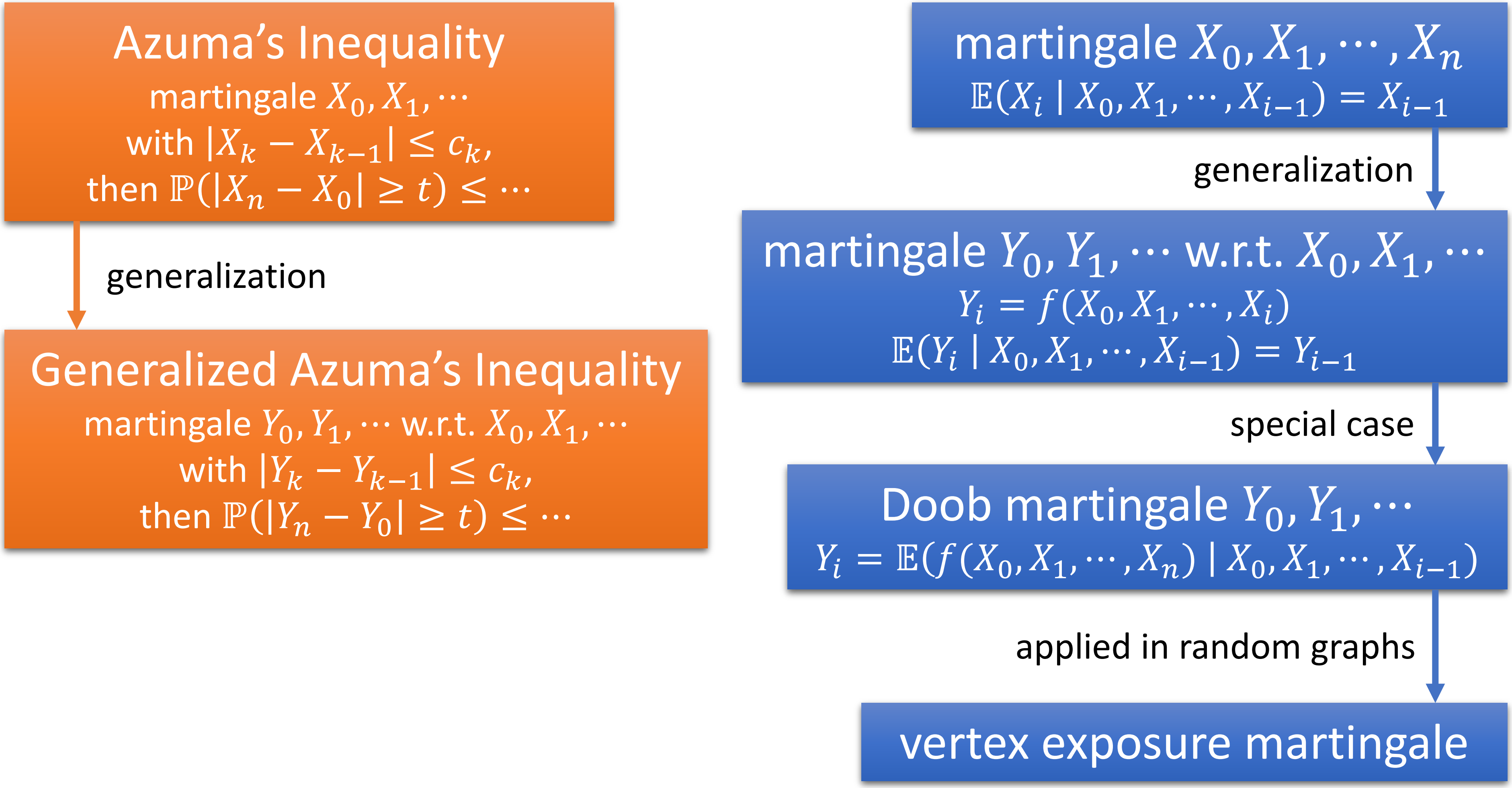
$Y_i = f(X_0, X_1, \cdots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

special case

Doob martingale $Y_0, Y_1, \cdots$

$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$



Azuma's Inequality
martingale $X_0, X_1, \cdots$
with $|X_k - X_{k-1}| \le c_k$,
then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$
generalization
Generalized Azuma's Inequality
martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
with $|Y_k - Y_{k-1}| \le c_k$,
then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$
martingale $X_0, X_1, \cdots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$
generalization
martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
$Y_i = f(X_0, X_1, \cdots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$
special case
Doob martingale $Y_0, Y_1, \cdots$
$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$
applied in random graphs
vertex exposure martingale

Azuma's Inequality

martingale $X_0, X_1, \cdots$

with $|X_k - X_{k-1}| \le c_k$,

then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

with $|Y_k - Y_{k-1}| \le c_k$,

then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

$Y_i = f(X_0, X_1, \cdots, X_i)$

$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

special case

Doob martingale $Y_0, Y_1, \cdots$

$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$

applied in random graphs

vertex exposure martingale

Sample Application:

Tight Concentration of Chromatic number

Doob Martingale + Generalized Azuma's Inequality

- for a function of (potentially dependent) r.v.:

$$f(X_1, X_2, \cdots, X_n)$$

- define corresponding Doob martingale:

$$Y_i = \mathbb{E}(f(X_1, \cdots, X_n) \mid X_1, \cdots, X_i)$$

  in particular, $Y_0 = \mathbb{E}(f(X_1, \cdots, X_n))$ and $Y_n = f(X_1, \cdots, X_n)$

- as long as the differences $|Y_i - Y_{i-1}|$ are bounded

- generalized Azuma's inequality implies $|Y_n - Y_0|$ is bounded

$\boxed{f(X_1, \cdots, X_n) \text{ is tightly concentration to its expectation}}$

The Method of Averaged Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \le i \le n$,

$$|\,\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})\,| \le c_i$$

Then,

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

The Method of Averaged Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \le i \le n$,

$$\left| \underbrace{\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i)}_{Y_i} - \underbrace{\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})}_{Y_{i-1}} \right| \le c_i$$

Then,

$$\mathbb{P}(|\underbrace{f(\boldsymbol{X})}_{Y_n} - \underbrace{\mathbb{E}(f(\boldsymbol{X}))}_{Y_0}| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Doob Martingale

The Method of Averaged Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \le i \le n$,

$$\left| \underbrace{\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i)}_{Y_i} - \underbrace{\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})}_{Y_{i-1}} \right| \le c_i$$

Then,

$$\mathbb{P}(|\underbrace{f(\boldsymbol{X})}_{Y_n} - \underbrace{\mathbb{E}(f(\boldsymbol{X}))}_{Y_0}| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Doob Martingale + Generalized Azuma's Inequality

The Method of Averaged Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$| \, \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1}) \, | \leq c_i$$

Then,

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

The Method of Averaged Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\,\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})\,| \leq c_i$$

Then,

May be hard to check!

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

$$\left| \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1}) \right| \leq c_i$$

Lipschitz Condition

$$\cancel{\left| \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1}) \right| \leq c_i}$$

A function $f(X_1, \cdots, X_n)$ satisfies the *Lipschitz condition* with constants c_i where $1 \leq i \leq n$, if

$$\mid f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, x_n) -$$
$$f(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, x_n) \mid \leq c_i$$

Average-case:

$$|\,\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})\,| \le c_i$$

Worst-case:

A function $f(X_1, \cdots, X_n)$ satisfies the *Lipschitz condition* with constants c_i where $1 \le i \le n$, if

$$|\, f(x_1, \cdots, x_{i-1}, x_i, x_{i+1}, x_n) -$$
$$f(x_1, \cdots, x_{i-1}, y_i, x_{i+1}, x_n)\,| \le c_i$$

The Method of Bounded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be n independent random variables and let $f(\boldsymbol{X})$ be a function satisfying the Lipschitz condition with constants c_i where $1 \le i \le n$, then:

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^{n} c_i^2}\right)$$

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\,\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})\,| \leq c_i$$

Then,

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

ded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be n independent random variables and let $f(\boldsymbol{X})$ be a function satisfying the Lipschitz condition with constants c_i where $1 \leq i \leq n$, then:

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \leq i \leq n$,

$$|\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})| \leq c_i$$

Then,

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

ded Differences

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be n independent random variables and let $f(\boldsymbol{X})$ be a function satisfying the Lipschitz condition with constants c_i where $1 \leq i \leq n$, then:

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \geq t) \leq 2 \exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Lipschitz condition

\+

Independence

$\longrightarrow$

bounded averaged differences

$$|\mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_{i-1})| \leq c_i$$

Azuma's Inequality

martingale $X_0, X_1, \cdots$
with $|X_k - X_{k-1}| \le c_k$,
then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
with $|Y_k - Y_{k-1}| \le c_k$,
then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

$Y_i = f(X_0, X_1, \cdots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

special case

Doob martingale $Y_0, Y_1, \cdots$

$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$

Azuma's Inequality

martingale $X_0, X_1, \cdots$
with $|X_k - X_{k-1}| \le c_k$,
then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$

generalization

Generalized Azuma's Inequality

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
with $|Y_k - Y_{k-1}| \le c_k$,
then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$

martingale $X_0, X_1, \cdots, X_n$

$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$

generalization

martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$

$Y_i = f(X_0, X_1, \cdots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$

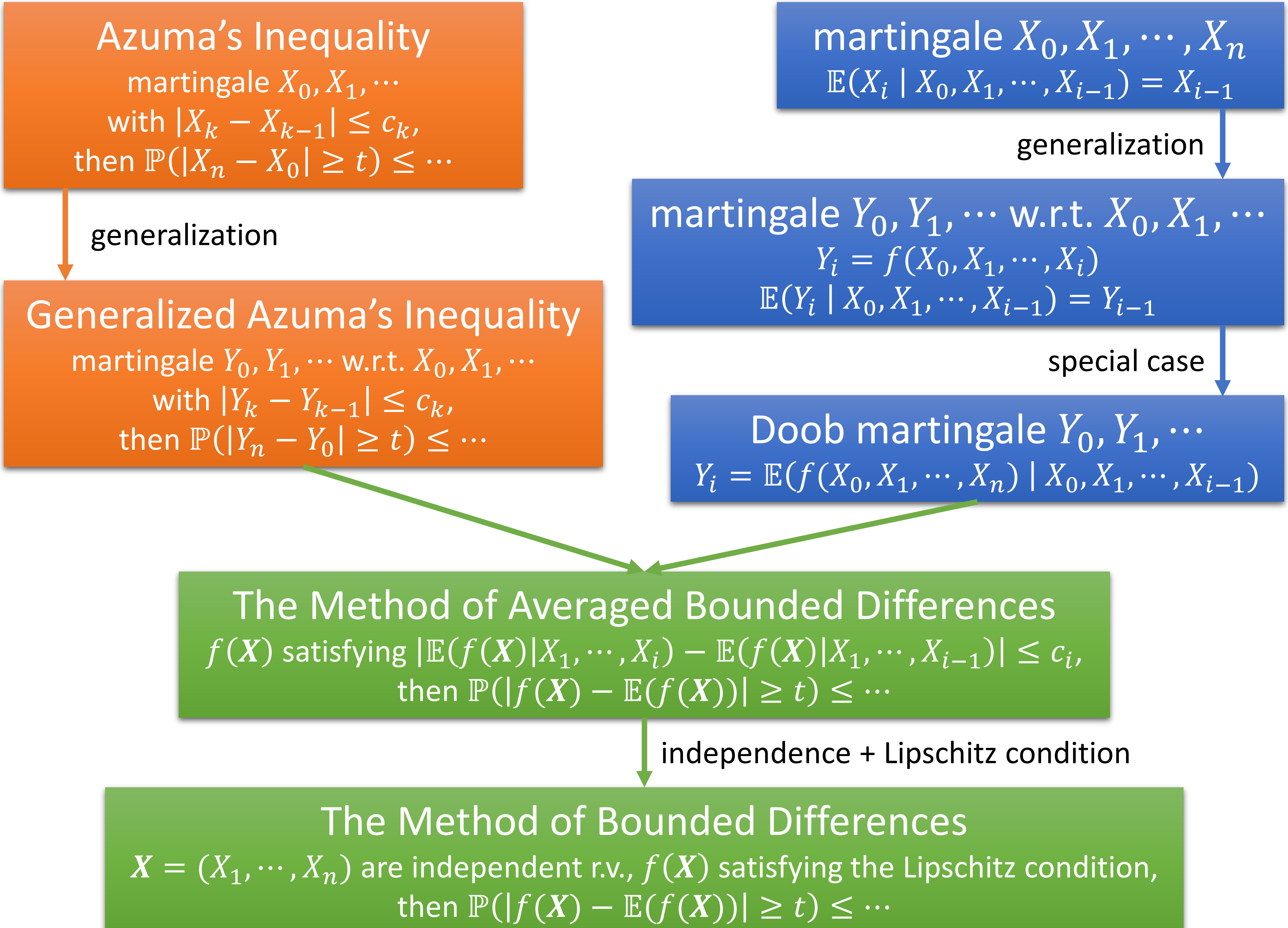
special case

Doob martingale $Y_0, Y_1, \cdots$

$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$

The Method of Averaged Bounded Differences

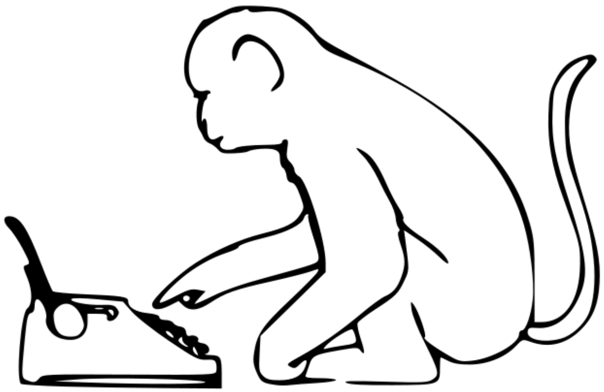
$f(\boldsymbol{X})$ satisfying $|\mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_{i-1})| \le c_i$,
then $\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le \cdots$


Azuma's Inequality
martingale $X_0, X_1, \cdots$
with $|X_k - X_{k-1}| \le c_k$,
then $\mathbb{P}(|X_n - X_0| \ge t) \le \cdots$
generalization
Generalized Azuma's Inequality
martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
with $|Y_k - Y_{k-1}| \le c_k$,
then $\mathbb{P}(|Y_n - Y_0| \ge t) \le \cdots$
martingale $X_0, X_1, \cdots, X_n$
$\mathbb{E}(X_i \mid X_0, X_1, \cdots, X_{i-1}) = X_{i-1}$
generalization
martingale $Y_0, Y_1, \cdots$ w.r.t. $X_0, X_1, \cdots$
$Y_i = f(X_0, X_1, \cdots, X_i)$
$\mathbb{E}(Y_i \mid X_0, X_1, \cdots, X_{i-1}) = Y_{i-1}$
special case
Doob martingale $Y_0, Y_1, \cdots$
$Y_i = \mathbb{E}(f(X_0, X_1, \cdots, X_n) \mid X_0, X_1, \cdots, X_{i-1})$
The Method of Averaged Bounded Differences
$f(\boldsymbol{X})$ satisfying $|\mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X})|X_1, \cdots, X_{i-1})| \le c_i$,
then $\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le \cdots$
independence + Lipschitz condition
The Method of Bounded Differences
$\boldsymbol{X} = (X_1, \cdots, X_n)$ are independent r.v., $f(\boldsymbol{X})$ satisfying the Lipschitz condition,
then $\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le \cdots$

The Method of Bounded Differences in Action:

Pattern Matching

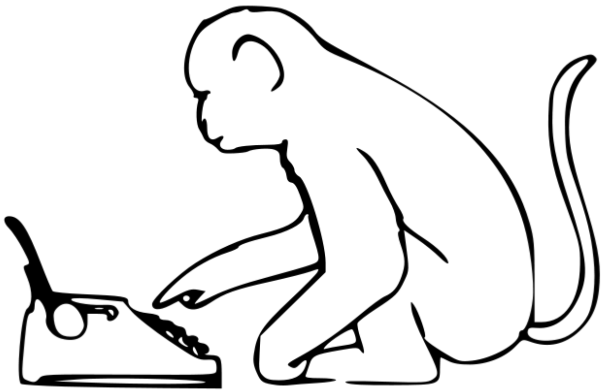
- a random string of length n
- a pattern of length k
- # of matched substrings?



The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?

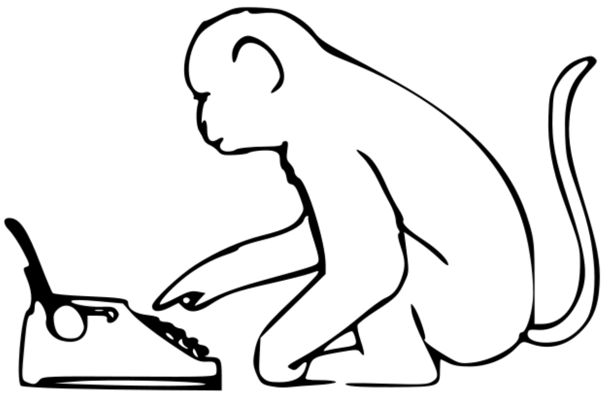


an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?



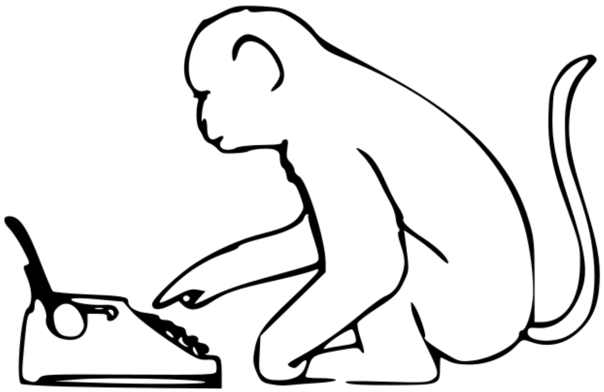
an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

The Method of Bounded Differences in Action:
Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?



an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

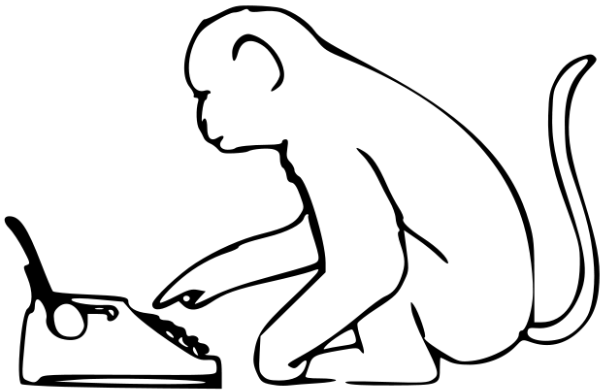
independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

$$\mathbb{E}(Y) = (n - k + 1)\left(\frac{1}{m}\right)^k$$

The Method of Bounded Differences in Action: Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?



an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

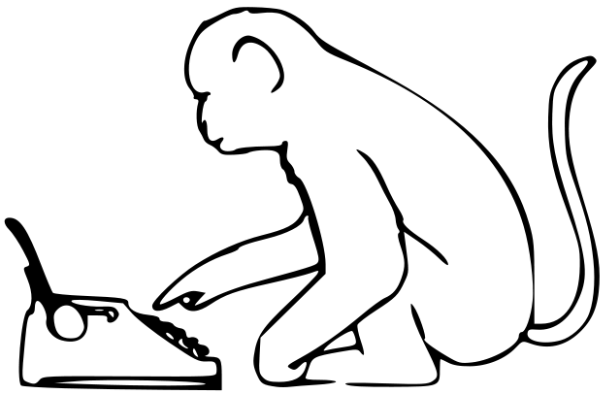
$$\mathbb{E}(Y) = (n-k+1)\left(\frac{1}{m}\right)^k$$

Deviation?

The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?



an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

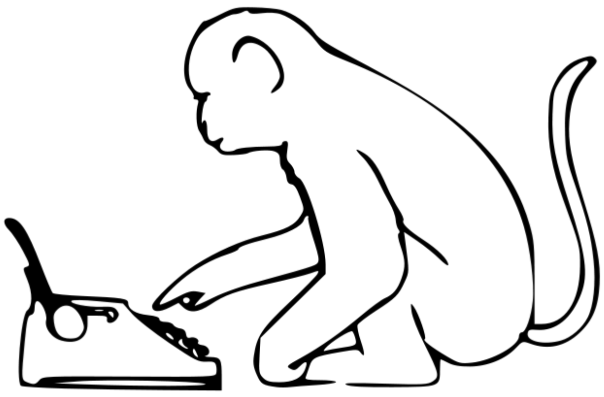
independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?



an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

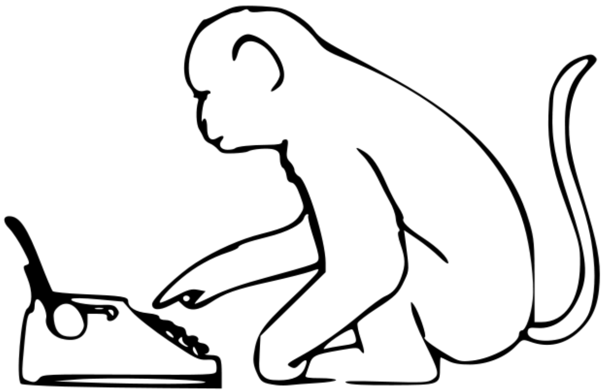
independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

$Y = f(X_1, X_2, \cdots, X_n)$

The Method of Bounded Differences in Action: Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?

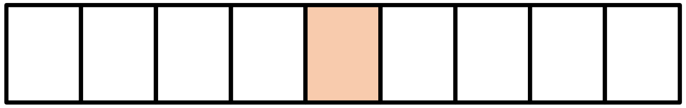


an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

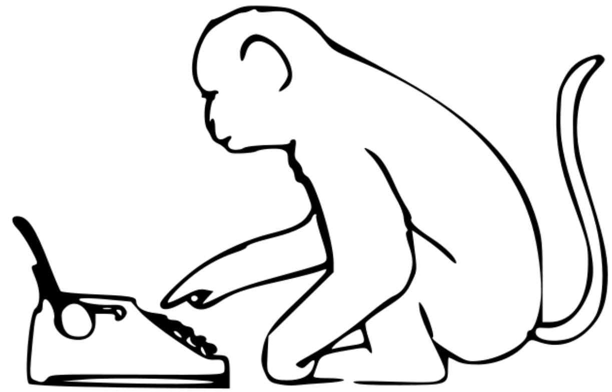
$Y = f(X_1, X_2, \cdots, X_n)$



The Method of Bounded Differences in Action:

Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?

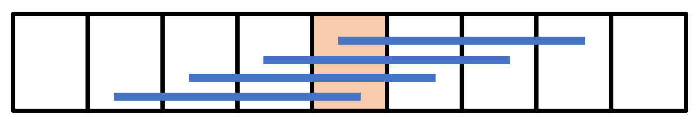


an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

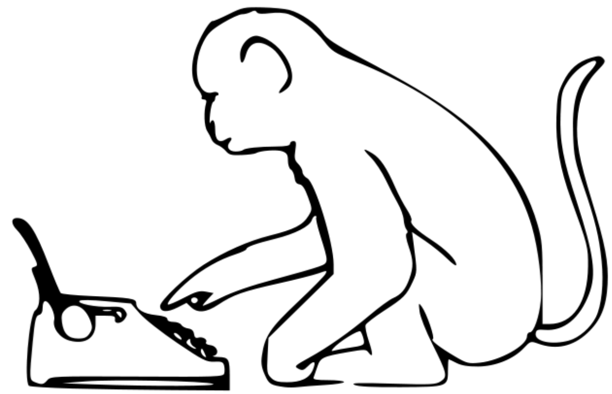
let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

$Y = f(X_1, X_2, \cdots, X_n)$



The Method of Bounded Differences in Action:
Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?

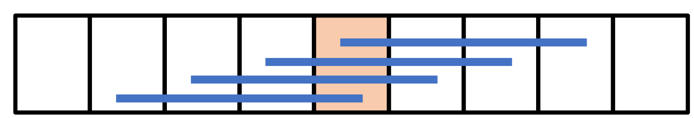


an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

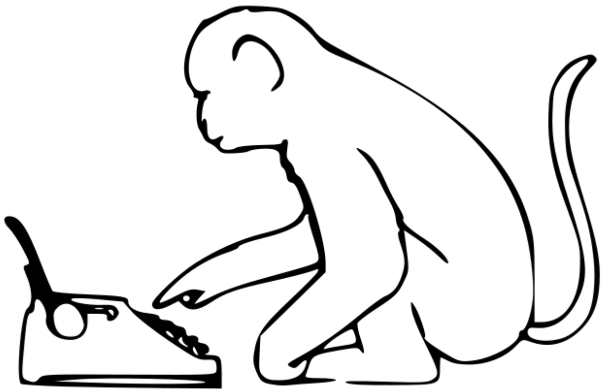
$Y = f(X_1, X_2, \cdots, X_n)$



changing any X_i changes f for at most k

The Method of Bounded Differences in Action:
Pattern Matching

- a random string of length n
- a pattern of length k
- # of matched substrings?

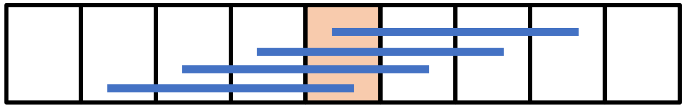


an alphabet Σ with $|\Sigma| = m$, a fixed pattern $\pi \in \Sigma^k$

independently and uniformly generate: $X_1, X_2, \cdots, X_n \in \Sigma$

let Y be number of substrings π in $\langle X_1, X_2, \cdots, X_n \rangle$

$Y = f(X_1, X_2, \cdots, X_n)$



changing any X_i changes f for at most k

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq tk\sqrt{n}) \leq 2e^{-t^2/2}$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$,
let $X = \sum_{i=1}^{n} X_i$, and $\mu = \mathbb{E}(X)$, then:

for any $\delta > 0$,

$$\mathbb{P}(X \geq (1+\delta)\mu) \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^{\mu}$$

for $0 < \delta < 1$,

$$\mathbb{P}(X \leq (1-\delta)\mu) \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^{\mu}$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0,1\}$,
let X

for an

for 0

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let $X = \sum_{i=1}^n X_i$, then:

for any $t > 0$,

$$\mathbb{P}(X \geq \mathbb{E}(X) + t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

$$\mathbb{P}(X \leq \mathbb{E}(X) - t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0,1\}$,
let X

for an

for 0

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let X

for an

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$
such that for all $k \geq 1$,

$$|Y_k - Y_{k-1}| \leq c_k$$

Then,

$$\mathbb{P}(|Y_n - Y_0| \geq t) \leq 2\exp\left(-\frac{t^2}{2\sum_{k=1}^{n} c_k^2}\right)$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0,1\}$,
let X

for an

for 0

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let X

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$ satisfying that, for all $1 \le i \le n$,

$$|\,\mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_i) - \mathbb{E}(f(\boldsymbol{X}) \mid X_1, \cdots, X_{i-1})\,| \le c_i$$

Then,

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0,1\}$,
let X

for an

for 0

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let X

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$
satisfying that, for all $1 \le i \le n$,

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be n independent random variables and let $f(\boldsymbol{X})$ be a function satisfying the Lipschitz condition with constants c_i where $1 \le i \le n$, then:

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

Concentration Inequalities

Question: probability that X deviates more than δ from expectation?

For independent r.v. $X_1, X_2, \cdots, X_n \in \{0, 1\}$,
let X
for an
for 0

For independent r.v. $X_1, X_2, \cdots, X_n$ where $X_i \in [a_i, b_i]$,
let X

Let $Y_0, Y_1, \cdots$ be a martingale with respect to $X_0, X_1, \cdots$

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ and let f be a function of $X_0, X_1, \cdots, X_n$
satisfying that, for all $1 \le i \le n$,

Let $\boldsymbol{X} = (X_1, \cdots, X_n)$ be n independent random variables and let $f(\boldsymbol{X})$ be a function satisfying the Lipschitz condition with constants c_i where $1 \le i \le n$, then:

$$\mathbb{P}(|f(\boldsymbol{X}) - \mathbb{E}(f(\boldsymbol{X}))| \ge t) \le 2\exp\left(-\frac{t^2}{2\sum_{i=1}^n c_i^2}\right)$$

There are more of them!